UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/629,486 | 07/29/2003 | Jan I. Ben | Ben 2-16-1-10 | 7452 |

46363 7590 03/30/2010
WALL & TONG, LLP/
ALCATEL-LUCENT USA INC.
595 SHREWSBURY AVENUE
SHREWSBURY, NJ 07702

| EXAMINER | |
|---|---|
| SAINT CYR, LEONARD | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 2626 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 03/30/2010 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

**BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES**

Application Number: 10/629,486
Filing Date: July 29, 2003
Appellant(s): BEN ET AL.

________________
Ben et al
<u>For Appellant</u>

**EXAMINER'S ANSWER**

This is in response to the appeal brief filed 12/10/09 appealing from the Office action mailed 08/24/09

**(1) Real Party in Interest**

A statement identifying by name the real party in interest is contained in the brief.

**(2) Related Appeals and Interferences**

The examiner is not aware of any related appeals, interferences, or judicial proceedings which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

**(3) Status of Claims**

The statement of the status of claims contained in the brief is correct.

**(4) Status of Amendments After Final**

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

**(5) Summary of Claimed Subject Matter**

The summary of claimed subject matter contained in the brief is correct.

**(6) Grounds of Rejection to be Reviewed on Appeal**

The appellant's statement of the grounds of rejection to be reviewed on appeal is correct.

**(7) Claims Appendix**

The copy of the appealed claims contained in the Appendix to the brief is correct.

**(8) Evidence Relied Upon**

| | | |
|---|---|---|
| 7,065,416 | Weare et al | 6-2006 |
| 5,615,302 | McEachern | 3-1997 |
| 6,633,845 | Logan et al | 10-2003 |

**(9) Grounds of Rejection**

The following ground(s) of rejection are applicable to the appealed claims:

***Claim Rejections - 35 USC § 103***

1. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

Claims 1 – 19, 21 – 37, and 41 - 45 are rejected under 35 U.S.C. 103(a) as being unpatentable over Weare et al., (US Patent 7,065,416) in view of McEachern (US Patent 5,615,302), and further in view of Logan et al., (US patent 6,633,845).

Regarding claims 1, 24, 34, Weare et al. discloses a method for program content identification (see col. 6, lines 22-27), said method comprising the steps of:

for each of at least two media program subsets, performing the steps of (col.5, lines 15 – 22):

filtering each first frequency domain representation of blocks of said media program subset using a plurality of filters to develop a respective second frequency domain representation of each of said blocks of said media said second frequency domain representation of each of said blocks having a reduced number of frequency

coefficients with respect to said first frequency domain representation program (Performing FFT on the frame data is considered as the first frequency domain; the frame data after the critical band filter is considered as the second frequency domain; and the critical band filters represent the plurality of filters. The number of frequency coefficients in the second frequency domain is reduced, since after critical band filters, the bandwidth becomes smaller, so fewer frequency coefficients are required to represent the frame data at the same resolution; col.16, lines 39 – 48; col.28, lines 9 - 11).

However, Weare et al., do not specifically teach that said plurality of filters have center frequencies logarithmically spaced apart from each other with a logarithmic additive factor of 1/12; grouping frequency coefficients of said second frequency domain representation of said blocks to form segments and selecting a plurality of said segments; comparing selected segments to features of stored programs to identify thereby said media program subset; determining whether said subsequent media program subset exhibits similarities to said initial media program subset.

McEachern teaches this 1/12 octave filter center frequency spacing results in logarithmically spaced filters that are very closely centered at the frequencies of the linearly spaces harmonics (col.12, line 66 – col.13, line 2).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to use logarithm filters as taught by McEachern in Weare et al., because that provide a superior speech information extractor that functions in a

manner similar to the functioning of the human auditory system and possesses similar acoustical performance (col.2, line 60 – col.3, line 5).

However, Weare et al., in view of McEachern do not specifically grouping frequency coefficients of said second frequency domain representation of said blocks to form segments and selecting a plurality of said segments; comparing selected segments to features of stored programs to identify thereby said media program subset; determining whether said subsequent media program subset exhibits similarities to said initial media program subset.

Logan et al., teach that the feature vectors corresponding to **the sequence of frames are organized into segments**. For example, **contiguous sequences of feature vectors may be combined into corresponding segments that are each of 1 second duration.** The distortion between various segments of the song is measured in order to identify those segments that can be considered to the same and those that are dissimilar. By identifying those segments of the audio input that share similar cepstral features, the system has been able to automatically decipher the song's structure (Utilizing segments of sizes other than 1 second duration suggests storing at least 30 minutes worth of segments, since in multimedia applications, such as television programs, a longer segment duration is required to identify a media entity, because of the overall length duration of certain TV programs; col.5, lines 4 – 35, col.6, lines 53 – 56).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to group sequence of frames in segments as taught by

Logan et al., in Weare et al., in view of McEachern, because that would help identify specific songs (col.2, line 5).

Regarding claim 2, Logan et al. further disclose that each grouping of frequency coefficients of said second frequency domain to form a segment represents blocks that are consecutive in time in said media program ("sequence of frames"; col.5, lines 5 – 35).

Regarding claim 3, Weare et al. in view of Logan et al., further disclose that said plurality of filters are arranged in a group that processes a block at a time, the portion of said second frequency domain representation produced by said group for each block forms a frame, and wherein at least two frames are grouped to form a segment (Weare et al., see col. 18, Logan et al. col.5, lines 5 – 35).

Regarding claim 4, Logan et al., further disclose that said selected segments correspond to portions of said media program that are not contiguous in time (col.6, lines 60 – 62).

As per claim 5, Logan et al., further disclose that said plurality of filters includes at least a set of triangular filters (col.4, lines 39 – 47).

As per claim 6, Logan et al., further disclose that said plurality of includes at least a set of log-spaced triangular filters (col.4, lines 39 – 47).

Regarding claim 7, Weare et al. further disclose that the segments selected in said selecting step are those that have largest minimum segment energy (see col. 18, lines 10-15).

Regarding claim 8, Weare et al. further disclose that the segments selected in said selecting step are selected in accordance with prescribed constraints (see col. 18, line 66 - col. 19 line 2, where only selecting peaks that last for more than specified number of frames prevents the peaks from being too close).

Regarding claim 9, Logan et al., further suggest that the segments selected in said selecting step are selected for portions of said media program that correspond in time to prescribed search windows that are separated by gaps ("assuming the frames are 25 ms long and overlap each other by 12.5ms"; col.5, lines 5 – 12).

Regarding claim 10, Weare et al. further disclose that the segments selected in said selecting step are those that result in the selected segments having a maximum entropy over the selected segments (see col. 18, lines 12- 15, where the most energetic peaks are chosen, thus choosing the most entropic peaks).

Regarding claims 11- 13, Weare et al. further suggest that the step of normalizing said frequency coefficients in said second frequency domain representation after performing said grouping step, said normalization being performed on a per-segment basis; wherein said normalization includes performing at least a preceding-time normalization; an L2 normalization ("normalizing the sum"; see col. 16, lines 3-6).

Regarding claim 14, Weare et al. further disclose that the step of storing said selected segments in a database in association with an identifier of said media program (see col. 7, lines 59-65, where music is stored in a database and for generating play lists thus an identifier must be associated with the stored data).

Regarding claim 15, Weare et al. further disclose that the step of storing in said database information indicating timing of said selected segments (see col. 9, lines 16-21, where classifying the tempo in the database indicates timing of media segment).

Regarding claim 16, Weare et al. further disclose that said first frequency domain representation of blocks of said media program is developed by the steps of: digitizing an audio representation of said media program to be stored in said database (see col. 16, lines 41-44); dividing the digitized audio representation into blocks of a prescribed number of samples (see col. 16, lines 41-44, where the audio representation is divided into frames); smoothing said blocks using a filter (see col. 16, lines 45-47); and

converting said smoothed blocks into the frequency domain, wherein said smoothed blocks are represented by frequency coefficients (see col. 16, lines 39- 41).

As per claim 17, Logan et al., further disclose a hamming window filter (col.4, lines 25 -27).

Regarding claim 18, Weare et al. further disclose that each of said smoothed blocks are converted into the frequency domain in said converting step using a Fast Fourier Transform (FFT) (see col. 16, lines 39-41 and col. 23, lines 52-54).

As per claim 19, Logan et al., further disclose converting step using a discrete cosine transform (col.4, line 49).

Regarding claims 21, and 37, Weare et al. discloses identification of content identification (see col. 6, lines 22-27), comprising:

for each of at least two media program subsets, performing the steps of (col.5, lines 15 – 22):

filtering each first frequency domain representation of blocks of said media program subset using a plurality of filters to develop a respective second frequency domain representation of each of said blocks of said media said second frequency domain representation of each of said blocks having a reduced number of frequency coefficients with respect to said first frequency domain representation program

(Performing FFT on the frame data is considered as the first frequency domain; the frame data after the critical band filter is considered as the second frequency domain; and the critical band filters represent the plurality of filters. The number of frequency coefficients in the second frequency domain is reduced, since after critical band filters, the bandwidth becomes smaller, so fewer frequency coefficients are required to represent the frame data at the same resolution; col.16, lines 39 – 48; col.28, lines 9 - 11).

However, Weare et al., do not specifically teach that said plurality of filters have center frequencies logarithmically spaced apart from each other with a logarithmic additive factor of 1/12; grouping frequency coefficients of said second frequency domain representation of said blocks to form segments; storing at least 30 minutes worth of segments; and selecting a plurality of said segments.

McEachern teaches this 1/12 octave filter center frequency spacing results in logarithmically spaced filters that are very closely centered at the frequencies of the linearly spaces harmonics (col.12, line 66 – col.13, line 2).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to use logarithm filters as taught by McEachern in Weare et al., because that provide a superior speech information extractor that functions in a manner similar to the functioning of the human auditory system and possesses similar acoustical performance (col.2, line 60 – col.3, line 5).

However, Weare et al., in view of McEachern do not specifically grouping frequency coefficients of said second frequency domain representation of said blocks to

form segments; storing at least 30 minutes worth of segments; and selecting a plurality of said segments.

Logan et al., teach that the feature vectors corresponding to **the sequence of frames are organized into segments**. For example, contiguous sequences of feature vectors may be combined into corresponding segments that are each of 1 second duration. **Obviously, segments of sizes other than 1 second may be utilized**. By identifying those segments of the audio input that share similar cepstral features, the system has been able to automatically decipher the song's structure (Utilizing segments of sizes other than 1 second duration suggests storing at least 30 minutes worth of segments, since in multimedia applications, such as television programs, a longer segment duration is required to identify a media entity, because of the overall length duration of certain TV programs; col.5, lines 4 – 35, col.6, lines 53 – 56).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to group sequence of frames in segments as taught by Logan et al., in Weare et al., in view of McEachern, because that would help identify specific songs (col.2, line 5).

As per claim 22, Weare et al., teach an apparatus for program content identification comprising:

a plurality of filters for filtering a first representation of a media program subset using frequency coefficient to develop a second representation of said media subset that has a reduced number of frequency coefficients with respect to said first

representation for each of at least two media program subsets (Performing FFT on the frame data is considered as the first frequency domain; the frame data after the critical band filter is considered as the second frequency domain; and the critical band filters represent the plurality of filters. The number of frequency coefficients in the second frequency domain is reduced, since after critical band filters, the bandwidth becomes smaller, so fewer frequency coefficients are required to represent the frame data at the same resolution; col.16, lines 39 – 48; col.28, lines 9 - 11).

However, Weare et al., do not specifically teach that said plurality of filters have center frequencies logarithmically spaced apart from each other with a logarithmic additive factor of 1/12; means for grouping ones of said coefficients of said second representation to form segments; means for storing at least 30 minutes worth of segments; and means for selecting a plurality of said segments.

McEachern teaches this 1/12 octave filter center frequency spacing results in logarithmically spaced filters that are very closely centered at the frequencies of the linearly spaces harmonics (col.12, line 66 – col.13, line 2).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to use logarithm filters as taught by McEachern in Weare et al., because that provide a superior speech information extractor that functions in a manner similar to the functioning of the human auditory system and possesses similar acoustical performance (col.2, line 60 – col.3, line 5).

However, Weare et al., in view of McEachern do not specifically means for grouping ones of said coefficients of said second representation to form segments;

means for storing at least 30 minutes worth of segments; and means for selecting a plurality of said segments.

Logan et al., teach that the feature vectors corresponding to **the sequence of frames are organized into segments**. For example, contiguous sequences of feature vectors may be combined into corresponding segments that are each of 1 second duration. Assuming the frames are 25 ms long and overlap each other by 12.5 ms, as described above, there will be approximately 80 feature vectors per segment. Obviously, **segments of sizes other than 1 second may be utilized**. By identifying those segments of the audio input that share similar cepstral features, the system has been able to automatically decipher the song's structure (Utilizing segments of sizes other than 1 second duration suggests storing at least 30 minutes worth of segments, since in multimedia applications, such as television programs, a longer segment duration is required to identify a media entity, because of the overall length duration of certain TV programs; col.5, lines 4 – 35, col.6, lines 53 – 56).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to group sequence of frames in segments as taught by Logan et al., in Weare et al., in view of McEachern, because that would help identify specific songs (col.2, line 5).

As per claim 23, Weare et al., teach an apparatus for program content identification comprising:

filtering a first frequency domain representation of a media program subset using a plurality of filters to develop a second frequency domain representation of each of said subsets of said media program having a reduced number of frequency coefficients with in said second frequency domain representation with respect to said first frequency domain representation for each of at least two media program subsets (Performing FFT on the frame data is considered as the first frequency domain; the frame data after the critical band filter is considered as the second frequency domain; and the critical band filters represent the plurality of filters. The number of frequency coefficients in the second frequency domain is reduced, since after critical band filters, the bandwidth becomes smaller, so fewer frequency coefficients are required to represent the frame data at the same resolution; col.16, lines 39 – 48; col.28, lines 9 - 11).

However, Weare et al., do not specifically teach means for filtering, wherein said plurality of filters have center frequencies logarithmically spaced apart from each other with a logarithmic additive factor of 1/12; means for grouping ones of said coefficients of said second representation to form segments; means for storing at least 30 minutes worth of segments; and means for selecting a plurality of said segments.

McEachern teaches this 1/12 octave filter center frequency spacing results in logarithmically spaced filters that are very closely centered at the frequencies of the linearly spaces harmonics (col.12, line 66 – col.13, line 2).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to use logarithm filters as taught by McEachern in Weare et al., because that provide a superior speech information extractor that functions in a

manner similar to the functioning of the human auditory system and possesses similar acoustical performance (col.2, line 60 – col.3, line 5).

However, Weare et al., in view of McEachern do not specifically means for grouping ones of said coefficients of said second representation to form segments; means for storing at least 30 minutes worth of segments; and means for selecting a plurality of said segments.

Logan et al., teach that the feature vectors corresponding to **the sequence of frames are organized into segments**. For example, contiguous sequences of feature vectors may be combined into corresponding segments that are each of 1 second duration. Assuming the frames are 25 ms long and overlap each other by 12.5 ms, as described above, there will be approximately 80 feature vectors per segment. Obviously, **segments of sizes other than 1 second may be utilized**. By identifying those segments of the audio input that share similar cepstral features, the system has been able to automatically decipher the song's structure (Utilizing segments of sizes other than 1 second duration suggests storing at least 30 minutes worth of segments, since in multimedia applications, such as television programs, a longer segment duration is required to identify a media entity, because of the overall length duration of certain TV programs; col.5, lines 4 – 35, col.6, lines 53 – 56).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to group sequence of frames in segments as taught by Logan et al., in Weare et al., in view of McEachern, because that would help identify specific songs (col.2, line 5).

Regarding claim 25, Weare et al., further disclose that the step of indicating that said media program cannot be identified when substantially matching segments are not found in said database in said searching step ("media entities that have...dissimilar"; Abstract).

Regarding claim 26, Logan et al., further disclose that said data base includes information indicating timing of segments of each respective media program identified therein, and wherein a match may be found in said searching step only when the timing of said segments produced in said grouping step substantially matches the timing of said segments stored in said database ("similar cepstral features, the system has been able to automatically decipher the song's structure"; col.6, lines 53 – 56).

Regarding claim 27, Weare et al., further disclose that said matching between segments is based on the Euclidean distances between segments (col.11, lines 15 – 20).

Regarding claim 28, Weare et al., further disclose that the step of identifying said media program as being the media program indicated by the identifier stored in said database having a best matching score when substantially matching segments are found in said database in said searching step ("matching algorithm...confidence level"; col.8, lines 1 - 12).

Regarding claim 29, Weare et al., further disclose that the step of determining a speed differential between said media program and a media program identified in said identifying step ("rate of speed"; col.23, lines 1 – 5).

Regarding claims 30, 32, and 33, Logan et al., in view of McEachern, and further in view of Weare et al., do not disclose wherein said matching score for a program P.sub.i is determined by

$$P_i = \frac{1}{z}\sum_{j=1}^{z} f(S'_{j=1} - S_j(P_i)).$$

wherein said determining step is based on an overlap score.

However, since Weare et al., teach nearest neighbor and/or other matching algorithms may be utilized to locate songs that are similar…a confidence level for song classification may also be returned (col.8, lines 1 – 10). One having ordinary skill in the at the time the invention was made would have found it obvious to use a matching score in Logan et al., in view of McEachern, and further in view of Weare et al., because that would help classify media entities (col.5, lines 7 – 12).

As per claim 35, Weare et al., teach an apparatus for program content identification comprising:

filtering a first frequency domain representation of a media program subset using a plurality of filters to develop a second frequency domain representation of each of said

subsets of said media program having a reduced number of frequency coefficients with in said second frequency domain representation with respect to said first frequency domain representation for each of at least two media program subsets (Performing FFT on the frame data is considered as the first frequency domain; the frame data after the critical band filter is considered as the second frequency domain; and the critical band filters represent the plurality of filters. The number of frequency coefficients in the second frequency domain is reduced, since after critical band filters, the bandwidth becomes smaller, so fewer frequency coefficients are required to represent the frame data at the same resolution; col.16, lines 39 – 48; col.28, lines 9 - 11).

However, Weare et al., do not specifically teach means for filtering, wherein said plurality of filters have center frequencies logarithmically spaced apart from each other with a logarithmic additive factor of 1/12; means for grouping ones of said coefficients of said second representation to form segments; means for searching a database for substantially matching segments, said database having stored therein segments of media programs and respective corresponding program identifiers; and means for determining whether said subsequent media program subset exhibits similarities to said initial media program subset.

McEachern teaches this 1/12 octave filter center frequency spacing results in logarithmically spaced filters that are very closely centered at the frequencies of the linearly spaces harmonics (col.12, line 66 – col.13, line 2).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to use logarithm filters as taught by McEachern in Weare

et al., because that provide a superior speech information extractor that functions in a manner similar to the functioning of the human auditory system and possesses similar acoustical performance (col.2, line 60 – col.3, line 5).

However, Weare et al., in view of McEachern do not specifically means for grouping ones of said coefficients of said second representation to form segments; means for searching a database for substantially matching segments, said database having stored therein segments of media programs and respective corresponding program identifiers; and means for determining whether said subsequent media program subset exhibits similarities to said initial media program subset.

Logan et al., teach that the feature vectors corresponding to **the sequence of frames are organized into segments**. For example, contiguous sequences of feature vectors may be combined into corresponding segments that are each of 1 second duration. Assuming the frames are 25 ms long and overlap each other by 12.5 ms, as described above, there will be approximately 80 feature vectors per segment. Obviously, **segments of sizes other than 1 second may be utilized**. By identifying those segments of the audio input that share similar cepstral features, the system has been able to automatically decipher the song's structure (Utilizing segments of sizes other than 1 second duration suggests storing at least 30 minutes worth of segments, since in multimedia applications, such as television programs, a longer segment duration is required to identify a media entity, because of the overall length duration of certain TV programs; col.5, lines 4 – 35, col.6, lines 53 – 56).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to group sequence of frames in segments as taught by Logan et al., in Weare et al., in view of McEachern, because that would help identify specific songs (col.2, line 5).

Regarding claim 36, Weare et al., in view of Logan et al., further disclose that said first frequency domain representation of said media program comprises a plurality of blocks of coefficients corresponding to respective time domain sections of said media program and said second frequency domain representation of said media program comprises a plurality of blocks of coefficients corresponding to respective time domain sections of said media program (Logan et al; col.5, lines 5 – 35; Weare et al., col.16, lines 33 – 36).

As per claims 41 – 45, Weare et al., further disclose at least two of said media subsets are associated with the same media program; at least two of said media subsets are associated with different media program ("media entities that are audio files or have portions that are audio files'; Abstract).

**(10) Response to Argument**

Appellants argue that neither Weare et al., nor McEachern nor Logan et al., teach or suggest filtering each frequency domain representation of blocks of a media program subset using a plurality of filters to develop a respective second frequency

domain representation of each of said blocks of said media program subset, said second frequency domain representation of each of said blocks having a reduced number of frequency coefficients with respect to said first frequency domain representation (Appeal Brief, pages 11 – 21).

The examiner disagrees, since Weare et al., disclose "A media entity is received by the system and **the data is converted from the time domain to the frequency domain via a Fast Fourier Transform (FFT). The FFT is performed on the frame data to produce a raw digital representation of the spectral characteristics of the media entity**. Subsequently, each frame may be processed in the following manner. **For each frame of data, at 750, critical band filtering is performed on the data, and the average of the data is calculated at 765**" (col.16, lines 39 – 48; col.28, lines 9 - 11). Performing FFT on the frame data is considered as the first frequency domain; the frame data after the critical band filter is considered as the second frequency domain; and the critical band filters represent the plurality of filters. The number of frequency coefficients in the second frequency domain is reduced, since after critical band filters, the bandwidth becomes smaller, so fewer frequency coefficients are required to represent the frame data at the same resolution.

Appellants argue that the motivation stated by the examiner to use logarithm filters fails to provide some articulated reasoning with some rational underpinning to support legal conclusion of obviousness (Appeal Brief, page 14).

The examiner points out that a new motivation of obviousness is given, which states that it is obvious to use logarithm filters because that provide a superior speech information extractor that functions in a manner similar to the functioning of the human auditory system and possesses similar acoustical performance (col.2, line 60 – col.3, line 5).

Appellants argue that neither Weare et al., nor McEachern nor Logan et al., storing at least 30 minutes worth of segments (Appeal Brief, pages 16 – 21).

The examiner disagrees, and points out that Logan et al., suggest that limitation by disclosing "the feature vectors corresponding to the sequence of frames are organized into segments...contiguous sequences of feature vectors may be combined into corresponding segments that are of 1 second duration...Obviously segments of sizes other than 1 second may be utilized" (col.5, lines 2 – 15; col.1, lines 8 -11). Utilizing segments of sizes other than 1 second duration suggests storing at least 30 minutes worth of segments, since in multimedia applications, such as television programs, a longer segment duration is required to identify a media entity, because of the overall length duration of certain TV programs.

Appellants argue that neither Weare et al., nor McEachern nor Logan et al., means searching a database for substantially matching segments, said database having stored therein segments of media programs and respective corresponding program identifiers (Appeal Brief, pages 17 - 21).

The examiner disagrees, since Logan et al., disclose "**By identifying those segments of the audio input** (e.g., the first half of the song being summarized) that **share similar cepstral features, the system has been able to automatically decipher the song's structure**" (col.6, lines 52 -57).

Appellants argue that the "means for" limitation recited in the invention cannot be broadly interpreted by the examiner to read on the implementation taught by Weare et al (Appeal Brief, page 18).

The examiner disagrees, since the "means for" limitations have been rejected over Logan et al., in view of McEachern. Please see claims rejection.

**(11) Related Proceeding(s) Appendix**

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

/Leonard Saint-Cyr/
Examiner, Art Unit 2626

/Richemond Dorvil/
Supervisory Patent Examiner, Art Unit 2626

**Conferees**:

Richmond Dorvil
/R. D./
Supervisory Patent Examiner, Art Unit 2626

Talivaldis Smits
/T. I. S./
Primary Examiner, Art Unit 2626